Last author version of:
Berendt, B. (2017). Better Data Protection by Design through multicriteria decision making: On false tradeoffs between privacy and utility. In E. Schweighofer, H. Leitold, A. Mitrakas, & K. Rannenberg (Eds.), GDPR & ePrivacy. APF 2017. Pre-Privacy Technologies and Policy. 5th Annual Privacy Forum, APF 2017, Vienna, Austria, June 7-8, 2017, Revised Selected Papers (pp. 210-230). Berlin etc.: Springer. LNCS 10518. (C) Springer.

Better Data Protection by Design through Multicriteria Decision Making: On False Tradeoffs between Privacy and Utility

Bettina Berendt

Dept. of Computer Science, KU Leuven, Leuven, Belgium,
firstname.lastname@cs.kuleuven.be,
https://people.cs.kuleuven.be/∼bettina.berendt/

Abstract. Data Protection by Design (DPbD, also known as Privacy by Design) has received much attention in recent years as a method for building data protection into IT systems from the start. In the EU, DPbD will become mandatory from 2018 onwards under the GDPR. In earlier work, we emphasized the multi*disciplinary* nature of DPbD. The present paper builds on this to argue that DPbD also needs a multi*criteria* approach that goes beyond the traditional focus on (data) privacy (even if understood in its multiple meanings).
The paper is based on the results of a survey (n=101) among employees of a large institution concerning the introduction of technology that tracks some of their behaviour. Even though a substantial portion of respondents are security/privacy researchers, concerns revolved strongly around social consequences of the technology change, usability issues, and transparency. The results taken together indicate that the decrease in privacy through data collection was associated with (a) an increase in accountability, (b) the blocking of non-authorized uses of resources, (c) a decrease in usability, (d) an altered perception of a communal space, (e) altered actions in the communal space, and (f) an increased salience of how decisions are made and communicated. These results call into question the models from computer science / data mining that posit a privacy-utility tradeoff. Instead, this paper argues, multicriteria notions of utility are needed, and this leads to design spaces in which less privacy may be associated with less utility rather than be compensated for by more utility, as the standard tradeoff models suggest. The paper concludes with an outlook on activities aimed at raising awareness and bringing the wider notion of DPbD into decision processes.

Keywords: Implementation aspects of "by design" and "by default" paradigms; Aspects of privacy impact and risk assessment; user studies; privacy and utility modelling and decision making

1 Introduction

Data Protection by Design (DPbD) has received much attention in recent years as an approach for building data protection into IT systems from the start. In the EU, DPbD will become mandatory from 2018 onwards under the General Data Protection Regulation (GDPR). In a panel at APF 2015, summarized and elaborated on in [21], we emphasized the need for multi*disciplinary* approaches to DPbD and illustrated this with conceptual and empirical examples.

A significant part of DPbD is the Data Protection Impact Assessment (DPIA) in which, among other things, the likely impacts of the planned technology on stakeholders' privacy are assessed. In multidisciplinary DPbD/DPIA, this notion of privacy (impacts) will be interpreted not only from a computational standpoint (where well-defined notions of data security and data confidentiality will be central), but based on a wider understanding of privacy including legal, sociological and psychological aspects. And in line with the GDPR requirement to implement appropriate technical and organisational measures to effect data protection, design solutions must be based on state-of-the-art methods for reducing unnecessary disclosures of personal data and/or avoiding unnecessary inference channels towards personal information.

The present paper reports on a survey that started from a multidisciplinary notion of privacy, but discovered in the answers a much richer set of concerns. This shows that DPbD also needs a multi*criteria* approach that goes beyond the traditional focus on data protection and privacy (even if understood in its multiple meanings).

In this context, it should be noted that the terms "data protection" and "privacy" are not defined uniformly and often used synonymously. Therefore, DPbD and DPIA are also often referred to as "Privacy by Design" and "Privacy Impact Assessment", e.g. [6, 5, 21]. In the survey described here, both terms were avoided. In the discussion, I will use the terms "data protection" in the sense of EU law and "privacy" in the general sense that "privacy can be violated by data processing", i.e. the (vague, but commonly used) concept at the intersection of the EU fundamental rights to privacy and to data protection. I return to a more differentiated notion of the two terms in the Conclusion.

The paper is organised as follows: Section 2 describes the case study and discusses its results. Section 3 investigates the implications, in particular of the exploratory analysis for risk-utility or privacy-utility tradeoffs in DPbD. Section 4 discusses limitations, and Section 5 summarises the conclusions, and gives an outlook on future work. Related work is referenced throughout the text to enhance readability.

2 A case study: Tracking coffee consumption

2.1 Context

Organisational and technological context The technology introduction took place in a computer-science department of a large research organization

that is offering free coffee, tea and mineral water to staff in a room open to all on the top floor of their building. For a number of years, two industrial coffee machines have been serving hot drinks. They are operated and filled with raw materials by an external service provider. (Hot water could be obtained from these machines or a separate, household-size electric kettle, with tea bags available on the table. Most tea drinkers use the electric kettle.) A paper sign on the door to the staff cafeteria, and paper signs on the coffee machines, communicate that room and drinks are for staff and their guests only. Doors are open during business hours and accessible via personnel-card readers at other times.

In November 2016, personnel-card readers were installed on the two coffee machines. The measure was announced in an email from the Head of Unit and explained further in a second email, in answer to a question by a staff member. The emails stated that the only change would be the need to swipe the card before getting coffee, and that "[t]here are no plans to collect statistics on everybody's consumption, neither the type of consumption nor the frequency". The costs of coffee per year were mentioned, and the benefits of the department only having to pay for its members' coffee consumption.

The personnel cards are contactless cards that are identified by a card number, which in turn is linked to the personnel ID. A card swipe causes an authentication request to the central authentication server. Upon successful authentication (= the cardholder is authorized to operate the resource), an electric circuit is closed for a few seconds and thus allows the resource to function (= the door to unblock, the coffee machine to dispense a hot drink, ...). All authentication requests are logged with card ID, resource ID, and timestamp.

The authentication server manages and logs data concerning the close to 70,000 members of the whole organisation. The department investigated in the present study employs 239 people, all of whom are authorised to consume coffee from the machines studied.

The cards grant access to a number of resources (originally doors and other gates) and have to be presented to a card reader to get food in the staff canteens. The recent deployment on coffee machines is in line with plans for employing the cards to control access to a wide range of other services (more doors, cupboards containing expensive equipment, ...). The organisation favours the deployment, for new purposes, of this form of access control because it is cheap (every member already has a card, the readers are cheap, and the authentication technology is in place anyway).

DPbD considerations The collected data are personal data, since they encode that a person identifiable via their card ID to their personnel ID, operated a certain resource at a certain time. (The card ID acts as a pseudonym.) The logged data encode the location and the fact of usage, i.e. *that* a drink was taken. Are they also sensitive data, i.e. special categories of data as per EU Data Protection Directive and GDPR? Food and drink choices are generally considered not to be sensitive data, although they may allow for inferences towards sensitive data, as

the example of food preferences on airline flights shows[1] [13]. Other inferences are possible, for example from food/drink consumption to health status. In the current example, the evidence for this is weaker, first because "coffee consumption" and even "coffee addiction" are generally not considered true health risks. Also, the card reader authentication does not lead to the logging of *what* drink the identifiable user chose (although some respondents believe that, see below); it could have been hot water or chocolate. Another possible type of inference is that the fact of using the coffee machine signals that the employee takes a break, but these inferences too are uncertain, since drinks are often taken specifically to accompany periods of intensive work, meetings, etc.

Motion profiles appear to be the most problematic type of inferences. With the planned increases in roll-out of card-reader authentication, these profiles can become increasingly fine-grained.

The introduction of such technology constitutes a potential application case for DPbD, since such data collection and processing falls under data protection law. The Article 29 Working Party observed [1, p.4, original emphasis] "Data protection requirements apply to the monitoring and surveillance of workers whether in terms of email use, Internet access, video cameras or location data. *Any monitoring must be a proportionate* response by an employer to the risks it faces taking into account the legitimate privacy and other interests of workers. *Any personal data* held or used in the course of monitoring must be *adequate, relevant and not excessive for the purpose for which the monitoring is justified.* Any monitoring must be carried out in the least intrusive way possible." The possibility of relying on the employer's "legitimate interest" is emphasized in [7], and a more general regulatory analysis is given in [15].

A good case study. Additional factors made this a good case study: "Coffee tracking", while a clear case of tracking, is perceived as neutral or even amusing by a majority of employees (as opposed to, say, a tracking of physician visits would be), and there is in general a high level of trust in the organisation and its data-protection integrity. This opened up a space in which people felt free to voice their concerns.

2.2 Research questions

Staff members reacted in different ways to the introduction of the new technology, ranging from simply accepting the measure and bringing and swiping their cards to get coffee, via short discussions of possible reasons, to extended "hacker-humour" discussions of how to break the system. On the whole, it appeared that questions and dissatisfaction persisted even after a number of weeks, and this situation motivated the research.

The research was, in part, motivated by our work on Privacy/Data protection by design [21]. In addition, the question was what this change in technology and

[1] For example, halal/kosher food preference can be an indicator of religion.

user interface of the coffee machines meant for privacy-related decision making in human-computer interaction.

Specifically, questions revolved around (1) the prerequisites for people to use PETs, in particular knowledge, beliefs and attitudes, in addition to or even before usability can improve technology acceptance, as posited in the step model of [18], and (2) the importance of social influences (as opposed to purely individual criteria such as cost-benefit analysis) on whether and how to use PETs, as posited by the ASPECT/ARCADE model of [17]. I was also interested in whether the knowledge about social effects can be leveraged for PETs technically via notions such as co-utility [10] and techniques such as collaborative distributed anonymisation [20], but decided to not foreground this complex construct.

The research questions were as follows:

- **(RQ1)** To what extent is the intent to use PETs dependent on prior knowledge of the underlying data collection technology?
- **(RQ2)** What are privacy-related decisions based on? In particular, to what extent is decision-making individual, and to what extent subject to social influences?
- **(RQ3)** Co-design: The new access control method constituted one design option. Would employees (as one stakeholder group) be able to generate more design options, and what would characterise these choices?

An additional motivation was to get an impression of attitudes and thoughts among employees, in order to understand the underlying current of discontent. This led to the practical questions

- **(PQ1)** Did employees know what personal data were collected, and did they care?
- **(PQ2)** Would they utilize an anonymization PET if it was available?

2.3 Method

A survey was created consisting of the following questions, all of which except (4) were open questions with free-form text answers (see Fig. 1 for an example):

- **(Q1)** Which data do you think are collected and stored by the card readers on the coffee machines?
- **(Q2)** Do you think the purpose of barring unauthorized coffee-getting could also be attained with other data, or other means? Please explain briefly.
- **(Q3)** If there was a button "anonymised version of coffee-getting authentication" on the card reader, would you use it? Feel free to make any assumptions about the system and its technology, as long as you briefly explain these assumptions.[2]
- **(Q4)** Do you think your use of the coffee machines (including your use of the fictitious button of question 3) may influence how your colleagues use them? (choice between "no", "maybe", "yes")

[2] The button was a fictitious PET, initially thought of as a version of distributed and possibly collaborative anonymisation [20], but open to interpretation by respondents.

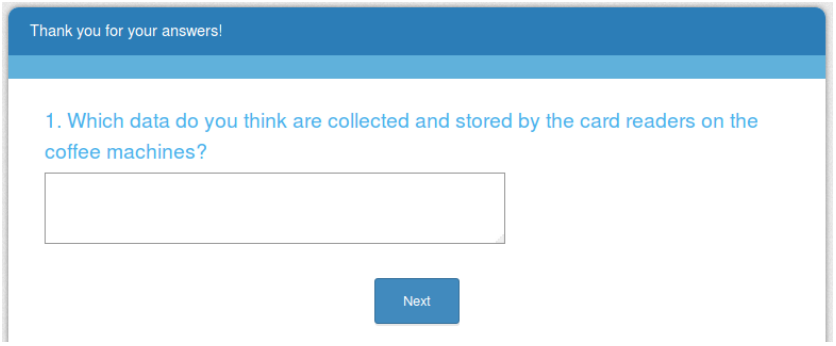

Fig. 1. Screenshot of the first survey question.

- **(Q5)** If you checked "maybe" or "yes" in the previous question, please explain why. Feel free to make any assumptions about the system and its technology, as long as you briefly explain these assumptions.
- **(comments)** Would you like to make any other comments?

The survey was available (a) online at `www.surveymonkey.com` and (b) as an RTF file that could be printed out and forwarded to me anonymously. The invitation to participate in the survey was sent out on 30th November, containing the link and the file, via the email alias that reaches all 239 people working in the department (i.e. all the people who may and now, via their personnel cards, can use the coffee machines). The survey was described as part of a research project. Answers were collected anonymously. To reduce the chance of people participating multiple times, the survey used the surveymonkey option that restricts answers to one per device. No further measures against multiple answers were taken; however, given everybodys time constraints it appeared unlikely that people would take part in a survey multiple times.

2.4 Results and discussion (1)

The survey was answered by 101 people, a response rate of 42%, within 2 days, with 1 paper and 100 online versions. One empty online result was not counted.[3] The free-form answers were analyzed with a simple form of thematic-analysis coding [3].

The large majority of answers were thoughtful and respectful, and respondents also expressed, in the comments question, much positive feedback about the fact of the study and interest in the results. Some answers were short and/or expressed humour and irony, such as "I'd like to suggest a James Bond-style iris-scan for coffee privileges. It would be way cooler and in that way I don't have

[3] It is possible, but irrelevant for the results, that this person had first flipped through the online version, was blocked from re-taking it, and therefore filled in the paper version.

to worry about forgetting my card." or "I am in a coffee drinking competition with a colleague. We want authenticated personal stats!"

The questions were asked in an order designed to constrain answers as little as possible, by asking a question q dealing with only one out of several possible to another question p, only after p. Q1 asked about what *knowledge (or beliefs)* people had about the new technology that had been deployed. Q2 asked about possible alternative technologies (or: *design* ideas). This question had to add a purpose (relative to Q1), without which the design question would have been meaningless, and it made a minimal assumption based on the official communication that had been made. Q3 could only be asked afterwards, since it asked about a specific instance of such technologies (anonymisation). Q4 was designed to explore an aspect that is characteristic of privacy and some PETs (including anonymisation PETs): social effects. With reference to PETs, it represented a further specialisation of Q3, and therefore had to be asked after it. Q5 was the request to elaborate on the multiple-choice answer to Q4.

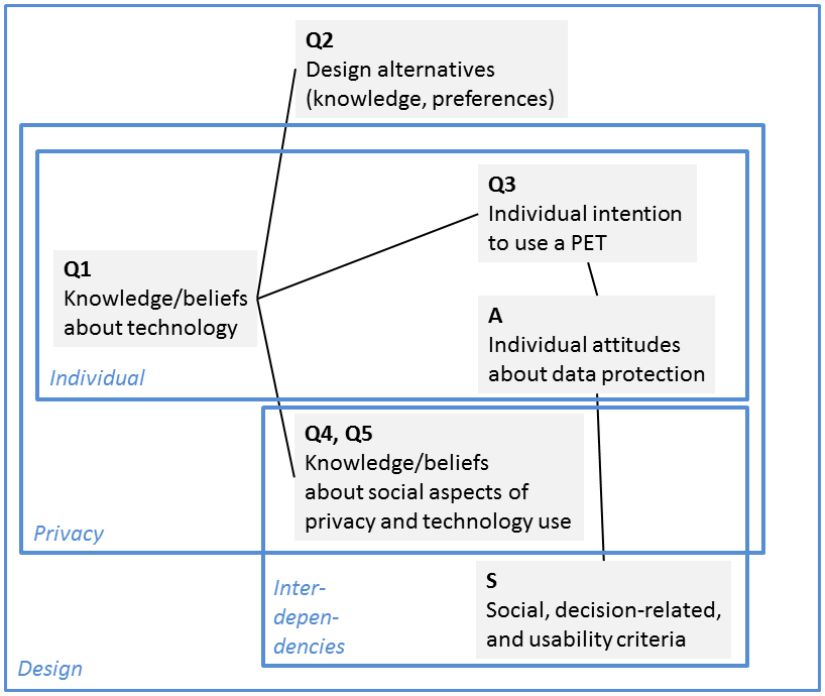

Fig. 2. Variables arising from the questions (Q1–Q5) and from the exploratory analysis (A, S), and their relationship to larger concepts. Lines between variables indicate the existence of dependencies.

For the purposes of discussing the results, a different order is more meaningful. Figure 2 shows a conceptual rendering of the decision aspects dealt with in the different questions. The analysis will cover variables describing individual knowledge/beliefs and intentions about technology and PETs, then variables that describe social knowledge/beliefs (and their intersection with privacy-related ones), and then design variables. A and S are variables defined in an exploratory phase of analysis (see Section 2.5).

Due to the open nature of the survey and the exploratory aspects of analysis, all data were analysed by descriptive statistics only. Inferential statistics are left for follow-up studies.

Q1. More than three quarters of respondents (rightly) believe that tracking takes place, but most do not know what is being tracked. 22.8% stated that no data were collected and stored. Only 26.7% gave the correct or near-correct answer that some user ID and the timestamp were being collected and stored. (Technically, only the card ID is logged, but of course this can be linked, via another database, to the user ID; therefore, these two answers were aggregated.) another 26.7% thought that also the type of drink was being logged. 9.9% considered that aggregate consumption only, maybe with something else, was recorded and stored, and 3% suspected some other form of data.

Answers including the type of drink would be correct if the respondent understood the question differently: "which data are collected and stored through a card swipe", because the coffee machine does record the type of consumed drink, even if neither the card reader nor the authentication server logs have access to this kind of information.

The distribution of answers may indicate that the question, while correctly targeting data-protection concerns, was not optimally phrased, since technically there are three different devices that collect data (coffee machine, card reader, authentication server), and all three have different methods and durations of storing the collected data. In addition, from a data-protection viewpoint, yet another question is crucial: whether the data are collected and/or stored and/or analysed. One respondent noted the logical underspecification in the information that had been communicated: "The promise that there are no plans to collect statistics does not say there will not be plans in the future, and statistics are not logging so one can later still create statistics of the past." Regardless of these different possible misinterpretations, it is worthwhile noting that nearly a quarter of employees, against their presumably existing understanding of how a personnel-card reader works, (wrongly) interpreted the "promise of no statistics" as "no data collection".

Q3 was answered in the affirmative by 36.6% of respondents, and in the negative by 35.6%. 12.9% mentioned usability ("one more button to press") as a factor, with 9 of these 13 people regarding the extra effort as a deterrent. (4 would still use it.) 10.9% mentioned trust ("how would I be able to check?", "although I would not trust it"), although 8 of these 11 people would still use it.

Q1–Q3. There was a clear connection between the beliefs about data collection and intentions to use the proposed PET (see Fig. 3). Obviously, for someone who believes that no data are being collected, it would make no sense to invest extra effort to anonymise, and thus the majority of these respondents (69.5%) would not use the button. (The four people who believed no data collection to take place in Q1 but wanting to use an anonymisation button in Q3 may have changed their beliefs along the survey). 43% of those who believe some kind of user ID is collected (whether with or without timestamp and with or without other information) intended to use the anonymisation button, and only 27.7% did not intend to use it. Most of the remaining 29% skipped the question.

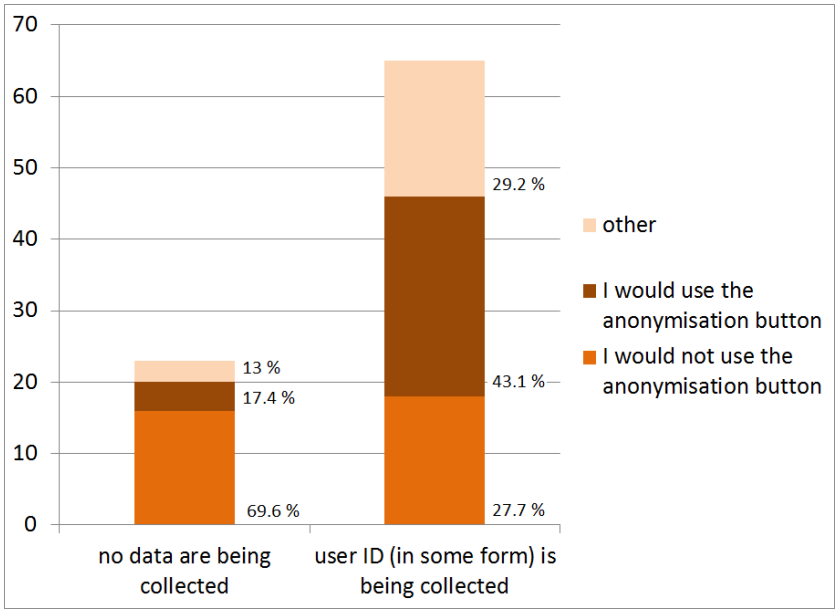

Fig. 3. Q1–Q3: knowledge/beliefs about data collection and intention to use the fictitious PET (numbers of respondents, percentages within the knowledge/belief group).

Q4 showed that nearly half of respondents (48.5%) believe that their use of the coffee machines has ("yes": 22.8%) or may have ("maybe": 25.7%) effects on other people. 34.7% explicitly believe that it does not have such effects ("no"). The remainder did not answer this question. The answers to question (5) showed that people had interpreted this question, as intended, with reference to coffee getting and data protection. Asked about the reasons why in question (5), only 1 person referred to k-anonymity (which had been the motivation for this question). 2 persons referred to effects of surveillance or chilling effects. Many more

talked more explicitly about social influence and peer pressure, with 7 referring to influence via imparting knowledge (awareness or reflection) and 9 via imitation.

Q1–Q4. Both Q1 and Q4 ask about knowledge or beliefs, one about a technology (not a PET), the other about privacy and behaviour in general and specific PETs in particular. One may therefore expect that the answer distributions should be independent of one another. On the other hand, it could be the case that people who are more knowledgeable or reflective about privacy in general might also be the ones for whom it is more salient that a particular technology will collect data. There is (weak) support for the latter hypothesis: A majority (50.7%) of those who believe some form of user ID is collected, also think that their behaviour influences others, whereas this percentage is only 43.4% among the "no data are collected" respondents.

Q2. Respondents proposed many different design alternatives. Since some people offered various alternatives and others offered none, the following percentages do not add up to 100%. The most frequent answers were two: there is no, or no efficient, alternative (18.8%), and social control (24.8%). These can be considered, in the light of data collection, as two ends of the spectrum: "the problem of unauthorised coffee-getting exists (and the current amount of data needs to be collected)" and "the problem does not exist or is negligible (and the previous approach in which no data are collected, is sufficient)". 15.8% explicitly said that they did not believe the problem exists.

Other answers acknowledge that the problem exists, but take different approaches with respect to data collection:

- **without data:** lock the room (4.9%), warning sign (4%), security guard (5%), no plastic cups and no cups in the cafeteria (1%)
- **without personal data (collection and/or storage):** anonymous tokens (10.9%), typing a code (6.9%), only checking authorization (5%)
- **with less, or less fine-grained, personal data:** restrict access to the cafeteria with card readers, at all hours (12.8%), or restrict access to the coffee machines with card readers, but only outside office hours (8.9%).

Three people mentioned cameras and facial recognition, i.e. more personal data, and two others suggested an interesting variant: *deactivated* card readers or cameras. Three answers suggested that people wanted to reap advantages of data collection (personalization of the drinks).

In sum, these answers suggest that co-design with the affected employees could work, and work efficiently (50% of respondents came up with their design ideas in less than 6.5min, and 79% in less than 15min). Of course, some of these options may in fact have been considered, and the relative costs are not known, but the abundance of answers belies the simple acceptance by those 18.8% who considered the chosen option to be without an alternative.

Q1–Q2. There was also evidence that the beliefs about data collection were associated with the activity level and type of co-design. Figure 4 shows the distribution of the design groups over the "no data collected" resp. "some form of user ID collected" respondents. Not only are data-collecting technologies that collect fewer or no personal data more popular among the latter; they also propose reliance on social control (as a specific form of "no data" design) very often. Conversely, most of the "no data collected" belief group sees no alternative to the current technology.

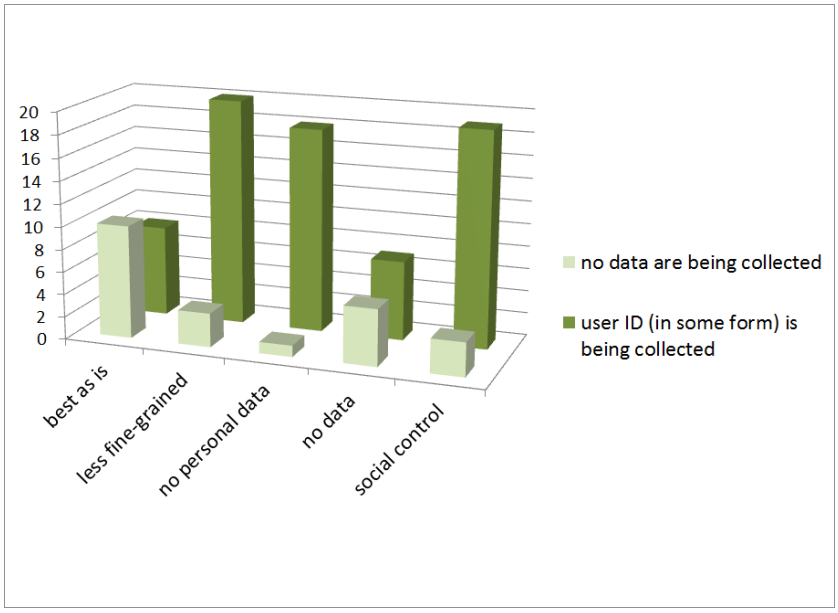

Fig. 4. Q1–Q2: knowledge/beliefs about data collection and proposals for design (numbers of respondents).

With regard to the research and practical questions, the following results were obtained:

- **(PQ1)** Most employees knew they were being tracked, but there were many misconceptions about details, and around a third believed there was no data collection/analysis.
- **(PQ2) and (RQ1)** Only slightly more than a third would use the fictitious PET, an equal number would not use it, and the rest did not answer. Knowledge/beliefs about data collection were clearly associated with intention to use.
- **(RQ2)** Nearly half of users believe there are social effects.

– **(RQ3)** Users came up with many alternative proposals in a short time, with people who had correct knowledge/beliefs about tracking generating more and more privacy-friendly variants.

2.5 Results and discussion (2): Exploratory analysis

When coding the free-form answers to Q2, Q3 and Q5, further recurring themes were identified.

A. Some respondents explicitly described their *privacy attitudes*, both positively (e.g., "a severe privacy infringement") and negatively (e.g., "I don't feel protective of this data in the slightest"). Privacy attitudes could also be inferred from many answers to the five questions, but attitudes per se were not inquired about, so only explicit mentions were coded for this new variable. The resulting variable A can therefore be regarded as a lower bound on the numbers of respondents with positive pro-privacy or negative non-privacy-concerned attitudes.

A–Q3. As could be expected, respondents who expressed pro-privacy attitudes also said they would use the fictitious PET (14 yes, 2 no), while respondents who expressed that they did not care about protecting these data showed the reverse pattern (4 yes, 12 no).

S. There were *usability* comments, both about the fictitious additional PET (see results of Q3 above) and about the existing technology (the card readers). Some remarks were made about *accountability*, via references to charging for coffee consumption. Most of these were negative, but some were neutral (coded as positive below).

In addition, respondents talked widely about effects that had not been expected in this research. First, these were *altered perceptions of the communal space*. This occurred both as a description of why the new technology was rejected (e.g., "The department always have felt like a place where everybody tries to be as flexible as possible [...] It's sad that the department now seems to be willing to question this flexibility over the price of a few coffees."), and as a design alternative (e.g. "provid[e] minimal free service to [other members of the organisaton], be kind and open to non personnel, ..."). Other comments described *altered actions in the communal space* (e.g., "Yes, [there will be social effects: employees] will buy own coffee machine for the office. This will reduce the number of informal meetings in cafeteria"). The events also led to an *increased salience of how decisions are made and communicated*, with most comments claiming that there had been no or poor communication of the purpose of the new technology, and no evidence given of non-authorised coffee-getting.[4]

All S mentions identified concerns, i.e. that respondents cared about these values. Thus, for S (sub-)variables, "positive" means the expression of a disutility

[4] In reaction to this, decision makers said that observed cases had in fact been communicated, and asked whether it was their task to prove abuse – which indeed would be impossible without another form of surveillance technology.

with respect to this criterion (e.g. for accountability: being charged for something that was previously free), whereas "negative" means the expression of a neutral or positive thought (e.g. that people would recognise the value of coffee).

The numbers of these comments are summarised in Table 1.

	A	usability (present technology)	usability (fictitious PET)	account-ability	altered perception	altered actions	saliency of decision-making	S total
positive	17	11	13	5	11	13	8	32
negative	18	0	0	2	1	1	0	4
sum	35	11	13	7	12	14	8	33

Table 1. Number of respondents with comments on A (explicitly expressed privacy attitudes) and on any of the other "social" criteria, summarised as S. Some totals of S are smaller than the column sum or row sum due to multiple criteria expressed by the same persons.

A–S. In total, 55 respondents (55.4%) made comments about A, S, or both. Of these, about 1/3 each (22 and 20) talked about only A or only S, and 13 about both. A further analysis of polarity indicated a substitution relationship between "privacy" and "social" rationales: Of those who had not commented on A, 30% commented on S. This proportion sank to 17.6% among those who had commented on A positively (i.e. expressed that privacy was important to them), but it rose to 55.6% among those who had commented on A negatively (i.e. expressed that privacy was not important to them). One respondent expressed this explicitly: "Anonymity is not the point here", then explaining their concern about S topics.

3 Consequences for risk/privacy-utility models in DPbD

The high response rate of the survey in general, and the free-form answers in particular, indicated that many employees perceived significant risks and disutility through the introduction of the card-reader access control. This has to be considered in relation to the utility gained.

A standard approach to this decision situation follows [11] and models

- **utility** is the utility of data usage.
- **risk** is the disclosure risk (or, more generally, privacy risk) to those whose personal data are being collected.

It is generally assumed that the processing of a full data set has the most utility, but also the most risk, the processing of no data has no utility and no risk, and fewer or transformed (e.g. k-anonymised) data have intermediate levels of

both. This produces the tradeoff "the more utility, the more risk". This is shown, schematically, by curve 1 in Fig. 5.

There is an alternative form of modelling, often used to describe and compare forms of privacy-preserving data mining/publishing. Here, the second component is a measure of (data) privacy, the inverse of risk, and the tradeoff is "the more privacy, the less utility" [2].

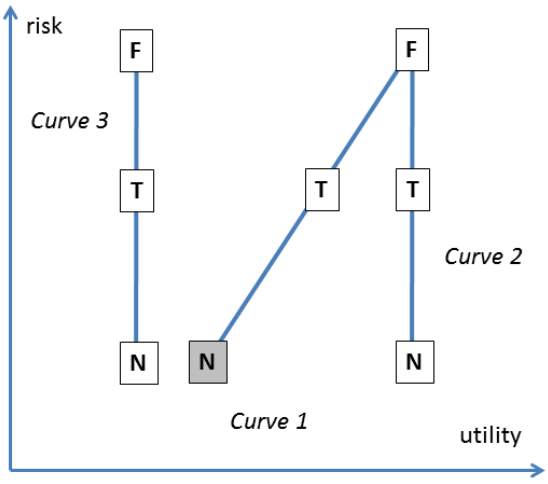

Fig. 5. Three schematic curves in risk-utility space. N = no personal data (with grey background: situation before start of data collection), F = full data, T = transformed data (e.g. k-anonymised).

The *existence* of this claimed tradeoff is, in a sense, tautological: If an unconstrained optimisation (e.g. the accuracy of a classifier learned from a full dataset) is the definition of the "full" utility, then any constrained optimisation (e.g. caused by some data privacy measure threshold) *must* be smaller or equal, and is usually smaller than, the unconstrained optimum. The *amount* that has to be traded off, and the *shape* of the tradeoff curve, may however depend on the data transformation processes applied.

Utility and privacy risk in the case study: a first model Going back to the case study: What exactly are utility and risk?

The utility in the original tradeoff curves is "data utility": "the value of a given data release as an analytical resource – the key issue being whether the data represent whatever it is they are supposed to represent" [12, p. 135]. However, this notion accords a conceptual independence to the data and their function that they do not have in real-world contexts and applications.

From the perspective of contexts and applications, utility is linked to the purpose of data collection. This is linked to different factors. If a given factor is part of the purpose, then a technology that achieves this goal will create utility. If it is not part of the purpose, it will not create utility. Here:

- **authentication** only personnel members should be able to use the resource.
- **accountability**$_1$ of coffee consumers for their consumption.
- **accountability**$_2$ of the coffee supplier for invoiced amounts.
- **accountability**$_3$ of individuals or specified anonymity sets in cases of abuse, as when a theft has occurred and the persons in the room at the time are to be determined.[5]

The privacy risks are mainly the potential to create motion profiles, make inferences from them, and take action based on these inferences (see Section 2.1).

3.1 Risk-utility tradeoffs, DPbD, and data minimisation

What would it mean to apply DPbD in the present use case, or in extensions of this? On the one hand, a number of concerns that could be termed "classical PET concerns" would need to be taken into account: the security and encryption technology used for cards, for data transfer, and log storage, access control for the logs, separation of the logs from the mapping of pseudonyms to IDs, etc. However, as Schaar [19] has pointed out in a critique of a classical case of the failure of an ambitious PbD project, a focus on these technology-centric considerations may lead decision makers to neglect important data protection principles, in particular data minimisation.

Data minimisation depends on the purpose of data collection. In a nutshell, it asks whether a given purpose can also be achieved with less data. This question presupposes that data processing yields a certain utility (= by fulfilling the purpose) and generates certain risks. Therefore, I propose to regard data minimisation as casting DPbD as a question of design in a risk-utility space.

For illustration, some simple approaches will be discussed that depart from the current infrastructure and thus would impose only negligible extra costs. The costs of card and authentication-server infrastructure are sunk costs; and the deployment of different hardware was not considered an alternative and is therefore not considered further here. To measure the extent of the risks, a simple variant of k-anonymity is used. (Different measures, e.g. based on differential privacy, or taking into account the accuracy of inferences, are possible but would require more assumptions.)

Assume the purpose is only to prevent unauthorised use of the resource and accountability$_2$. An anonymisation of the logs whereby each user pseudonym is replaced by the constant "authorised user" or "unauthorised user" would suffice to serve this purpose with *no personal data*, thus leaving utility unaffected and reducing privacy risk to their starting level. (For the sake of simplicity of

[5] In the case study, the latter two were mentioned by administrative/management personnel involved in card-reader deployment, in a follow-up interview of the survey.

the argument, threats from stronger – and more costly – attacks involving for example physical observation and record linkage, are ignored.)[6]

If the purpose is also accountability$_1$, a follow-up question needs to ask whether in fact individuals are to be charged for their consumptions, or administrative units. In the latter case, a k-anonymisation of the logs whereby each authorised user's pseudonym is replaced by their respective unit ID would suffice to serve this purpose. Again, utility would remain unaffected, and privacy risk would be reduced to the level of k, with k the size of the smallest unit.

The distinction between individual and collective accountability [4] becomes more acute if accountability$_3$ is also a purpose. Various questions should be asked: Should and could this accountability be individual or collective (for example, it is conceivable that whole units take responsibility and are held liable in cases of theft)? Should such accountability be multi-step, i.e. the unit takes responsibility to the outside and imposes individual sanctions on the inside? Who should decide on this question?

For all forms of accountability, data are likely to be needed only for certain periods. Beyond that, they can (and therefore should) be deleted, an operation that will not affect utility but reduce privacy risks.

Best-case risk-utility values resulting from this thought experiment are shown, in schematic form, by curve 2 in Fig. 5, which indicates that there is no or a negligible tradeoff (negligible if the costs of data transformation are taken into account, no if they aren't).

3.2 Extending risk-utility tradeoff models by multicriteria decision-making modelling

However, as the exploratory analysis has shown, there is a third component here summarised as S. In the present study, the following additional factors of (dis)utility were found (cf. Section 2.5):

- **usability**
- **altered perception of a communal space**
- **altered actions in the communal space**
- **increased salience of how decisions are made and communicated.**

This can be modelled as an additional risk factor or an additional disutility factor. All else equal, this would shift curves upward (more risk) or to the left (less utility). Curve 3 in Fig. 5 uses the latter approach in order to not change the semantics of the risk.[7] It shows that in extreme cases, even with perfect PETs

[6] This constraint on utility also illustrates the dependence of technical solutions' utility on purposes. The proposal "no plastic cups and no cups in the cafeteria" to Q2 would serve the purpose of barring non-authorised use, but not that of accountability$_2$. However, the existence of this purpose was likely unknown to respondents.

[7] The semantics of the risk that are generally used in risk-utility models focus on an individuum-centric notion of privacy. The current focus on the risks of tracking using personal data (see Section 3) follows this approach. Certainly privacy is not only an

and data minimisation, the outcome could be worse than the starting point: Assume that authentication is data-minimal and secure, and no personal data are stored. As a result, utility with respect to the goals of authentication and accountability may increase, as explained in the previous section. However, this increase may be more than offset by a decrease in utility caused by losses in usability and perceptions of and actions in the communcal space. The increased salience of decision making may be considered positive for utility (to the extent that employee awareness and participation are desired) or as negative for utility (to the extent that such awareness is considered to lead to discontent for employees and/or work for management). These utilities and disutilities may be experienced by different stakeholders, but they can be aggregated into an organisation-wide utility measure. In sum, *any* choice along curve 3, which contains the available options with the new technology, would be inferior to the starting point, i.e. create less utility and the same or higher risk.

To avoid such inferior choices, DPbD should draw more strongly on multicriteria decision making: Data protection and privacy risks need to be measured but should be weighed against a notion of utility composed of the classical purpose-dependent utility and disutilities caused by usability and social implications.

4 Limitations and lessons learned

There were only two clearly negative comments on the survey, and these serve well as introductions to this section.

One respondent found the questions "silly, not precise enough, and highly biased".

I believe that silliness is a matter of perspective and will therefore disregard this. As explained above, the questions were on purpose underspecified and left much room for open, including unexpected, answers. It is true, however, that this openness also in some cases led to answers that were more difficult to interpret. So while openness allowed for exploration and the discovery of the variables A and S, and thereby led to the design ideas described in Section 3.2, the results should be validated in follow-up work in a confirmatory manner.

There was indeed some imprecision in phrasing, in particular with respect to Q1. This issue has been described in Section 2.3, and as argued there, this imprecision also had some unexpected advantages. Still, in follow-up work a compromise should be found between technical exactness and linguistic simplicity when describing technological functionality.

In this study, there were two main expressions or sources of bias.

The first relates to Q2. It implied that the card readers have the purpose of barring unauthorised coffee-getting. In fact, this was an *interpretation* of the

individual but also a collective value, so some aspects of "altered perceptions of a communal space" could be modelled as an additional factor of privacy risk. However, it appears questionable to subsume also usability or the salience of decision making under "privacy risks".

communication to employees, which had not talked about purpose(s), but highlighted the benefit to the department of not having to pay for outsiders' coffee consumption (which had never been authorised, but previously could not be avoided). This interpretation led to the notion of "unauthorised coffee-getting" in Q2.

The phrasing of question (2) may suggest that this be the sole purpose. This was not the case (even if other purposes had not been communicated). While the phrasing of Q2 was legitimate in the context of Q2 (whose purpose was indeed to obtain design alternatives for this purpose), the phrasing may have influenced the answers to the questions following it. In follow-up work, the order of questions should be considered very carefully, and questions that may prime certain concepts may be placed later.

The second source of bias was made apparent by the results themselves, in particular in the exploratory analysis. The research and survey questions were formulated on a background of a long personal history of privacy research, and this may have led to a certain *déformation professionelle*. As Gürses and Diaz [14] observe, one always needs to ask who formulated the privacy problem: the "experts" or the "users". As they point out, privacy/security experts tend to perceive problems of institutional privacy, usually the collection and processing of data by powerful corporations or governments, whereas users tend to focus more on social privacy, the question of who among their peers should know what. In [9], we have proposed this question of "who defines the privacy problem" as one of five key self-reflective questions that privacy researchers should ask themselves to improve the quality and transparency of their work. The results of the present study suggest that part of the bias of the "expert" is, already prior to questions such as institutional or social privacy, to cast every problem as a privacy problem. While the present users' concerns about S topics often revolved around *social* consequences of technology, these were not limited, or even expressly not about, *social privacy*. When they were about privacy, they revolved around institutional privacy. Viewed over all users, there seemed to be a substitution effect of institutional privacy concerns versus social non-privacy concerns. Thus, the present study suggests an additional self-reflective question: "who defines the problem, and is it really (only) a privacy problem?".

The second critical remark from respondents was that "[t]his survey looks much like unnecessary criticism on the department's decision to install these card readers." As remarks from other respondents, referred to above as increased saliency of decision-making, show, many respondents voiced criticism of this decision. However, in the light of the De Hert and Gutwirth [8] analysis of data protection as a transparency tool towards the powerful (data controllers and processors)[8], it needs to be asked when such criticism is "unnecessary" and when it is not. Other respondents regarded the very existence of criticism as positive: "The critical reception of the card readers on our coffee machines is actually a good sign. One can't expect (junior) scientists to be good and uncritical at

[8] complemented by privacy protection as an opacity tool towards the powerless (data subjects)

the same time." The survey itself led to some concrete measures for improving the transparency of decision making (so far, a voluntary self-commitment of the employee representative to communicate results of decision making more widely).

5 Summary, general conclusions, and future work

The results of the case study validate models of complex individual decision-making in privacy-related questions, in particular the importance of social influences posited by the ASPECT/ARCADE model of [17]. They also validate the necessity of a number of prerequisites for people to use PETs, in particular knowledge, beliefs and attitudes, in addition to or even before usability can improve technology acceptance, as posited in the step model of [18].

The opinion, widespread among respondents, that privacy operates also via social effects can be interpreted, with some caution, as an understanding that "my privacy utility influences yours, and vice versa", thus presenting some empirical support for PETs that are built on co-utility [10]. However, concrete instances such as collaborative distributed anonymisation [20] require more dedicated user-based evaluations. In a first user study [16], we found that while non-technical users in general understood the concept of k-anonymity and the notion of privacy that it can provide for them, it was less clear to them that to obtain such k-anonymity, they need to contribute to it. Thus, the study did not provide evidence that these users understood or appreciated the notion of co-utility applicable in the example application and architecture. That study, however, suffered from an example domain in which participants did not have strong privacy preferences. Follow-up work will aim at designing a more convincing task, taking into account also the results of the present study.

Beyond supporting earlier research, the present results however call into question the models from computer science / data mining that posit a privacy-utility tradeoff, where utility is measured in a simplistic way that centers on the accuracy of the personal data and (if applicable) the models learned from these data. The results illustrate how social considerations, and considerations about – both their own and organisations' – decision-making and its transparency, are woven into people's reactions to technology. The perceived negative effects on social spaces can even outweigh perceived threats based on data processing and possible privacy violations. Therefore , multicriteria notions of utility are needed, and this leads to design spaces in which less privacy can be associated with less utility rather than be compensated for by more utility, as the standard tradeoff models suggest. From a legal standpoint, a multicriteria notion of utility already ties in well with the GDPRs stated goal that data protection be the protection of a wide range of individuals rights and freedoms, not only the rights to data protection and privacy. From a computational standpoint, however, more efforts are needed to embed multicriteria utility into DPbD.

In future work, we aim to use the insights gained for different phases of DPbD, in particular Impact Assessments and design itself. This includes creating practical guidelines for including these considerations into (a then extended)

Impact Assessment, and testing these guidelines for understandability and effectiveness. As a first step, this can build on the PIA Guidelines we developed for teaching and training contexts [21]. In addition, organisational card-reader deployment will be studied in a more general, recently started project involving the present author and others.

Acknowledgements I thank all respondents of the survey for their thought-inspiring answers, and all those involved in the "New Developments in data privacy" workshops 2016 for support and valuable ideas: the Cambridge University Isaac Newton Institute and Turing Gateway to Mathematics, the organisers Mark Elliot, Natalie Shlomo and Chris Skinner, and all participants. Ralf De Wolf has provided helpful comments an on earlier version of the text.

References

1. Article 29 Working Party (2001). *Opinion 8/2001 on the Processing of Personal Data in the Employment Context.* `http://ec.europa.eu/justice/data-protection/article-29/documentation/opinion-recommendation/files/2001/wp48_en.pdf` [2017-04-13]
2. Bertino, E., Lin, D., & Jiang, W. (2008). A survey of quantification of privacy preserving data mining algorithms. In C.C. Aggarwal & P.S. Yu (Eds.), *Privacy-preserving Data Mining: Models and Algorithms* (pp. 181-200). New York: Springer.
3. Boyatzis, R. (1998). *Transforming Qualitative Information: Thematic Analysis and Code Development.* London: Sage.
4. Bovens, M. (2006). *Analysing and Assessing Public Accountability. A Conceptual Framework.* European Governance Papers (EUROGOV) No. C-06-01. `http://www.connex-network.org/eurogov/pdf/egp-connex-C-06-01.pdf` [2017-04-13]
5. Crespo García, A. et al. (2016). *PRIPARE. Privacy- and Security-by design Methodology Handbook* `http://pripareproject.eu/wp-content/uploads/2013/11/PRIPARE-Methodology-Handbook-Final-Feb-24-2016.pdf` [2017-04-13]
6. Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Le Métayer, D., Tirtea, R., & Schiffner, S. (2014). *Privacy and Data Protection by Design from Policy to Engineering.* ENISA Report. `https://www.enisa.europa.eu/publications/privacy-and-data-protection-by-design` [2017-04-13]
7. Data Protection Commissioner (undated). *Guidance Note for Data Controllers on Location Data.* `https://www.dataprotection.ie/docs/Guidance-Note-for-Data-Controllers-on-Location-Data/1587.htm` [2017-04-13]
8. De Hert, P. and Gutwirth, S. (2006). Privacy, data protection and law enforcement. Opacity of the individual and transparency and power. In E. Claes, A. Duff, & S. Gutwirth (Eds.), *Privacy and the Criminal Law* (pp. 61-104). Antwerp/Oxford: Intersentia.
9. De Wolf, R., Vanderhoven, E., Berendt, B., Pierson, J., & Schellens, T. (2016). Self-reflection on privacy research in social networking sites. *Behaviour & Information Technology.* DOI: 10.1080/0144929X.2016.1242653.
10. Domingo-Ferrer, J., Martínez, S., Sánchez, D., & Soria-Comas, J. (2017). Co-Utility: Self-Enforcing protocols for the mutual benefit of participants. *Eng. Appl. of AI, 59,* 148-158.

11. Duncan, G.T., Keller-McNulty, S.A., & Stokes, S.L. (2001). *Disclosure Risk vs. Data Utility: The R-U Confidentiality Map.* National Institute of Statistical Sciences. Technical Report Number 121. `http://www.niss.org/sites/default/files/technicalreports/tr121.pdf` [2017-04-13]
12. Elliot, M., Mackey, E., O'Hary, K., & Tudor, C. (2016). *The Anonymisation Decision-Making Framework.* Manchester, UK: UKAN. `http://ukanon.net/wp-content/uploads/2015/05/The-Anonymisation-Decision-making-Framework.pdf` [2017-04-13]
13. European Union Agency For Fundamental Rights FRA (2014). *Twelve operational fundamental rights considerations for law enforcement when processing Passenger Name Record (PNR) data.* `https://fra.europa.eu/sites/default/files/fra-2014-fundamental-rights-considerations-pnr-data-en.pdf` [2017-04-13]
14. Gürses, S. & Diaz, C. (2013). Two tales of privacy in online social networks. *IEEE Security & Privacy, 11 (3)*, 2937.
15. Hendrickx, F. (2002). *Protection of Workers' Personal Data in the European Union: Two Studies.* `http://ec.europa.eu/social/BlobServlet/docId=2507` [2017-04-13]
16. Herelixka, E. (2016). *Experiencing a Privacy Enhancing Technology. An Exploratory User Study of Collaborative Anonymization.* Masters Thesis. KU Leuven, Faculty of Science.
17. Jameson, A., Berendt, B., Gabrielli, S., Cena, F., Gena, C., Vernero, F., & Reinecke, K. (2014). Choice architecture for Human-Computer Interaction. *Foundations and Trends in Human-Computer Interaction, 7(1-2),* 1-235.
18. Renaud, K., Volkamer, M., & Renkema-Padmos, A. (2014). Why doesn't Jane protect her privacy? In *Privacy Enhancing Technologies* (pp. 244-262).
19. Schaar, P. (2010). Privacy by Design. *Identity in the Information Society, 3(2),* 267-274.
20. Soria-Comas, J. & Domingo-Ferrer, J. (2015). Co-utile collaborative anonymization of microdata. In *Modeling Decisions for Artificial Intelligence* (pp. 192-206).
21. Tsormpatzoudi, P., Berendt, B., & Coudert, F. (2016). Privacy by Design: From research and policy to practice - the challenge of multi-disciplinarity. In *Privacy Technologies and Policy. Third Annual Privacy Forum, APF 2015. Luxembourg, Luxembourg, October 7-8, 2015. Revised Selected Papers* (pp. 199-212). Berlin etc.: Springer. LNCS 9484.